

SEQUENCE LISTING

(1) GENERAL INFORMATION:

-
- (i) APPLICANT: Dale, James B.
 - (ii) TITLE OF INVENTION: GROUP A STREPTOCOCCAL VACCINES
 - (iii) NUMBER OF SEQUENCES: 16
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SEED and BERRY LLP
 - (B) STREET: 6300 Columbia Center, 701 Fifth Avenue
 - (C) CITY: Seattle
 - (D) STATE: Washington
 - (E) COUNTRY: USA
 - (F) ZIP: 98104
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US
 - (B) FILING DATE: 12-SEP-1997
 - (C) CLASSIFICATION:
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: McMasters, David D.
 - (B) REGISTRATION NUMBER: 33,963
 - (C) REFERENCE/DOCKET NUMBER: 481112.410P1
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (206) 622-4900
 - (B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGGGGGCAT CGGTCGCGAC TAGGTCTCAG ACAGAT

36

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGGGGGGGAT CCACGTAGTT TCTCTTTAGC

30

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGGGGGGGAT CCGCCGTGAC TAGGGGTACA

30

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGGGGGTCG ACCTCAGTTT TTAACCCTTC

30

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGGGGGTCG ACAGAGTGTT TCCTAGGGGG

30

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGGGGCCAT GGTAACTTGT CATTATTAGC

30

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGGGGCCAT GGAGAGTGCG TTATACTAGG

30

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGGGGCTGC AGAGATAACT TCTCATTCTG

30

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGGGGCTGC AGAACGGTGA TGGTAATCCT

30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGGGGGGTA CCAGCTCTCT TAAAATCTCT

30

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGGGGGGTA CCTTGTTAGA TCAGGTTACA

30

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGGGGATCG ATATTAACT CTTGTAACAG

30

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGGGGGATCG ATGTCGCGAC TAGGTCTCAG

30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGGGGGAAGC TTTTACTTAC GTGCCTCTAA TTC

33

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1158 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCA TGC ATG GTC GCG ACT AGG TCT CAG ACA GAT ACT CTG GAA AAA GTA	48
Ala Cys Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val	
1 5 10 15	
CAA GAA CGT GCT GAC AAG TTT GAG ATA GAA AAC AAT ACG TTA AAA CTT	96
Gln Glu Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu	
20 25 30	
AAG AAT AGT GAC TTA AGT TTT AAT AAT AAA GCG TTA AAA GAT CAT AAT	144
Lys Asn Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn	
35 40 45	
GAT GAG TTA ACT GAA GAG TTG AGT AAT GCT AAA GAG AAA CTA CGT GGA	192
Asp Glu Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg Gly	
50 55 60	
TCC GCC GTG ACT AGG GGT ACA ATA AAT GAC CCG CAA AGA GCA AAA GAA	240
Ser Ala Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu	
65 70 75 80	
GCT CTT GAC AAG TAT GAG CTA GAA AAC CAT GAC TTA AAA ACT AAG AAT	288
Ala Leu Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Asn	
85 90 95	
GAA GGG TTA AAA ACT GAG AAT GAA GGG TTA AAA ACT GAG AAT GAA GGG	336
Glu Gly Leu Lys Thr Glu Asn Glu Gly Leu Lys Thr Glu Asn Glu Gly	
100 105 110	
TTA AAA ACT GAG AAT GAA GGG TTA AAA ACT GAG GTC GAC AGA GTG TTT	384
Leu Lys Thr Glu Asn Glu Gly Leu Lys Thr Glu Val Asp Arg Val Phe	
115 120 125	
CCT AGG GGG ACG GTA GAA AAC CCG GAC AAA GCA CGA GAA CTT CTT AAC	432
Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu Leu Asn	

130	135	140	
AAG TAT GAC GTC GAG AAC TCT ATG TTA CAA GCT AAT AAT GAC AAG TTA Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp Lys Leu			480
145	150	155	160
<hr/>			
165	170	175	
CCA TGG AGA GTG CGT TAT ACT AGG CAT ACG CCA GAA GAT AAG CTA AAA Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys Leu Lys			528
180	185	190	
<hr/>			
195	200	205	
AAT GAG AAG TTA TCT CTG CAG AAC GGT GAT GGT AAT CCT AGG GAA GTT Asn Glu Lys Leu Ser Leu Gln Asn Gly Asp Gly Asn Pro Arg Glu Val			624
210	215	220	
<hr/>			
225	230	235	
ATA GAA GAT CTT GCA GCA AAC AAT CCC GCA ATA CAA AAT ATA CGT TTA Ile Glu Asp Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu			672
240			
<hr/>			
245	250	255	
CGT CAC GAA AAC AAG GAC TTA AAA GCG AGA TTA GAG AAT GCA ATG GAA Arg His Glu Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu			720
260	265	270	
<hr/>			
275	280	285	
ACA CAA TTA TAT ACT AAA CAT AAT AGT AAT TAC CAA CAA TAT AAT GCA Thr Gln Leu Tyr Thr Lys His Asn Ser Asn Tyr Gln Gln Tyr Asn Ala			816
290	295	300	
<hr/>			
305	310	315	
GCG ACT AGG TCT CAG ACA GAT ACT CTG GAA AAA GTA CAA GAA CGT GCT Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala			960
320			

GAC AAG TTT GAG ATA GAA AAC AAT ACG TTA AAA CTT AAG AAT AGT GAC 1008
Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp
325 330 335

TTA AGT TTT AAT AAT AAA GCG TTA AAA GAT CAT AAT GAT GAG TTA ACT 1056
Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr
340 345 350

GAA GAG TTG AGT AAT GCT AAA GAG AAA CTA CGT AAA AAT GAT AAA TCA
Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg Lys Asn Asp Lys Ser
355 360 365

CTA TCT GAA AAA GCT AGT AAA ATT CAA GAA TTA GAG GCA CGT AAG
Leu Ser Glu Lys Ala Ser Lys Ile Gln Glu Leu Glu Ala Arg Lys
370 375 380

TAAAAGCTT 1158

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:

 - (A) LENGTH: 383 amino acids
 - (B) TYPE: amino acid
 - (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Cys Met Val Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val
 1 5 10 15

Gln Glu Arg Ala Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu
 20 25 30

Lys Asn Ser Asp Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn
 35 40 45

Asp Glu Leu Thr Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg Gly
 50 55 60

Ser Ala Val Thr Arg Gly Thr Ile Asn Asp Pro Gln Arg Ala Lys Glu
 65 70 75 80

Ala Leu Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Asn
 85 90 95

Glu Gly Leu Lys Thr Glu Asn Glu Gly Leu Lys Thr Glu Asn Glu Gly
 100 105 110

Leu Lys Thr Glu Asn Glu Gly Leu Lys Thr Glu Val Asp Arg Val Phe
 115 120 125

Pro Arg Gly Thr Val Glu Asn Pro Asp Lys Ala Arg Glu Leu Leu Asn
 130 135 140

Lys Tyr Asp Val Glu Asn Ser Met Leu Gln Ala Asn Asn Asp Lys Leu
 145 150 155 160

Pro Trp Arg Val Arg Tyr Thr Arg His Thr Pro Glu Asp Lys Leu Lys
 165 170 175

Lys Ile Ile Asp Asp Leu Asp Ala Lys Glu His Glu Leu Gln Gln Gln
 180 185 190

Asn Glu Lys Leu Ser Leu Gln Asn Gly Asp Gly Asn Pro Arg Glu Val
 195 200 205

Ile Glu Asp Leu Ala Ala Asn Asn Pro Ala Ile Gln Asn Ile Arg Leu
 210 215 220

Arg His Glu Asn Lys Asp Leu Lys Ala Arg Leu Glu Asn Ala Met Glu
 225 230 240

Val Ala Gly Arg Asp Phe Lys Arg Ala Gly Thr Leu Leu Asp Gln Val
 245 250 255

Thr Gln Leu Tyr Thr Lys His Asn Ser Asn Tyr Gln Gln Tyr Asn Ala
 260 265 270

Gln Ala Gly Arg Leu Asp Leu Arg Gln Lys Ala Glu Tyr Leu Lys Gly
 275 280 285

Leu Asn Asp Trp Ala Glu Arg Leu Leu Gln Glu Leu Asn Ile Asp Val
 290 295 300

Ala Thr Arg Ser Gln Thr Asp Thr Leu Glu Lys Val Gln Glu Arg Ala
 305 310 315 320

Asp Lys Phe Glu Ile Glu Asn Asn Thr Leu Lys Leu Lys Asn Ser Asp
325 330 335

Leu Ser Phe Asn Asn Lys Ala Leu Lys Asp His Asn Asp Glu Leu Thr
340 345 350

Glu Glu Leu Ser Asn Ala Lys Glu Lys Leu Arg Lys Asn Asp Lys Ser
355 360 365

Leu Ser Glu Lys Ala Ser Lys Ile Gln Glu Leu Glu Ala Arg Lys
370 375 380